NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

MEMORANDUM
RM-3044-PR
MAY 1962

POINT ESTIMATION OF RELIABILITY FROM RESULTS OF A SMALL NUMBER OF TRIALS

Z. A. Typaldos and D. E. Brimley

PREPARED FOR:

UNITED STATES AIR FORCE PROJECT RAND

The RAND Corporation
SANTA MONICA • CALIFORNIA

MEMORANDUM
RM-3044-PR
MAY 1962

POINT ESTIMATION OF RELIABILITY FROM RESULTS OF A SMALL NUMBER OF TRIALS

Z. A. Typaldos and D. E. Brimley

This research is sponsored by the United States Air Force under Project RAND — Contract No. AF 49(638)-700 — monitored by the Directorate of Development Planning, Deputy Chief of Staff, Research and Technology, Hq USAF. Views or conclusions contained in this Memorandum should not be interpreted as representing the official opinion or policy of the United States Air Force. Permission to quote from or reproduce portions of this Memorandum must be obtained from The RAND Corporation.

The RAND Corporation
1700 MAIN ST • SANTA MONICA • CALIFORNIA

PREFACE

This Memorandum is a part of continuing RAND studies on methods for improving the reliability of Air Force systems. Early estimation of reliability on the basis of a few trials is of value in that it focuses attention on critical reliability problems. This Memorandum provides a new approach to the problem of reliability estimation.

SUMMARY

This Memorandum presents a new approach to the estimation of reliability on the basis of a few trials. In this new method, based on information theory concepts, when the probability of success is not known, the proposed way to assign probabilities is to maximize the measure of uncertainty. This point-estimation method is characterized by less-extreme fluctuations in estimates as the results of Bernoulli trials become known.

CONTENTS

PREFACE iii

SUMMARY v

Section

I. INTRODUCTION 1

II. HEURISTIC DEVELOPMENT 3

III. EXAMPLES 8

IV. CONCLUSIONS 11

REFERENCES 13

I. INTRODUCTION

The process of testing to determine reliability is generally expensive and time-consuming. Under the pressures of costs and schedules, engineers are in many instances required to make estimates of reliability based on the results of tests of very few units. In the case of Bernoulli trials which have an outcome of either success or failure, the reliability is usually estimated as the number of successes divided by the total number of trials. Very little confidence can be placed in this method of estimation if the number of trials is small. Consider an extreme example: if a conventional coin has been tossed only once, the result yields an estimate for obtaining heads as either 1 or 0. These values are as far removed from the *a priori* value of 0.5 as it is possible to be on the probability scale. This type of experiment is, therefore, quite meaningless.

When possible outcomes of a test are success and failure, the familiar binomial distribution is applied to the solution of many problems. The equation for this distribution is

$$(p + q)^N = p^N + N\,p^{N-1}\,q + \ldots \frac{N!}{m!n!}\,p^m\,q^n \ldots + q^N = 1$$

where

p = probability of success

$q = 1 - p$ = probability of failure

N = number of trials

m = number of successes

$n = N - m$ = number of failures

Each term of the expansion represents the probability of obtaining exactly m successes and n failures in N trials. It will be observed that the value

of p must be assigned or estimated on the basis of previous knowledge of the characteristics of the population, or estimating as $\frac{m}{N}$. Many useful applications of this distribution exist when a priori knowledge yields the exact value of p, such as problems which are analogous to coin or die tossing or drawing cards from a pack. Difficulties in applications are in the estimation of p on the basis of a very few trials if previous knowledge of p is nonexistent. This problem has been addressed in the past by many mathematicians.

A brief summary of the past work on the problem is included in a paper by Steinhaus.[1] Generally, it is assumed that there is an a priori probability distribution, $f(p)$, for the unknown parameter p. After taking a sample, x, we then have an a posteriori estimate of $g(p;x)$ for p. The value of p is then assigned by applying one of several criteria, such as maximizing g, the mean value of $g(p;x)$ dp, or some other function, depending on whether we want the "most probable" value of p, the value that minimizes the mean square error, or a value that satisfies some other criterion. Baye's original tentative suggestion was to take a uniform a priori distribution, i.e., $f(p) = 1$, if one has no initial knowledge. More recently the work of Wald has suggested that other choices of f may be appropriate in various circumstances. Steinhaus discusses the situation when one assumes that there is a loss function proportional to the mean square error, and the loss function is to be minimized.

This Memorandum provides an approach to the problem based on information-theory concepts.

II. HEURISTIC DEVELOPMENT

An approach to the problem of estimating success probabilities on the knowledge of the outcome of a few trials can be based on the concepts of information theory developed by E. C. Shannon in Ref. 2, which is summarized briefly in Chapter 3 of Ref. 3. These concepts lead to the equation

$$H = -\sum_i P_i \ln P_i$$

where

H = measure of uncertainty

P_i = probability of the i^{th} state

$\ln P_i$ = a measure of the uncertainty in the i^{th} state

If we are concerned only with the future events of success and failure, there are only two possible states, and this equation reduces to

$$H = -p \ln p - q \ln q$$

If we have conducted a number of trials in which the outcome has been determined, the probability of m successes and n failures in N trials is given by the usual binomial distribution result for Bernoulli trials

$$P(m, n) = \frac{N!}{m!n!} p^m q^n$$

The logarithm of $P(m, n)$ is a measure of the information that has been obtained from a given number of trials, and since the information is in hand, the probability of obtaining it is 1. This reduces the uncertainty, so that we can write as a measure of the residual uncertainty

$$H = -p \ln p - q \ln q + 1(\ln \frac{N!}{m!n!} p^m q^n)$$

Here again if p is known from previous information, we have a straightforward calculation to determine the measure of the residual uncertainty H.

If p is not known, the proposed way to assign the probabilities is to maximize the function H. This maximizing principle, which can only be accepted as an axiom, has been used in other fields in which information theory is applied. In statistical mechanics, where the quantity analogous to the measure of uncertainty is the entropy, Jaynes[4,5] and Tribus[6] have successfully used the maximizing principle in the solution of problems. For example, Tribus obtains the probability distribution of the energy states of a thermodynamic system by considering these states as the possible outcomes of the system, with unknown probabilities of occurrence, and then maximizing the entropy of the distribution subject to the condition that the expected value of the energy of the system is known. The distribution thus obtained is that of the well-known "microcanonical ensemble" of statistical mechanics. This distribution is obtained more conventionally by a lengthy and difficult train of reasoning which is eliminated by the maximum-entropy principle.

If H is maximized, we have

$$+\ln \frac{p}{q} = \frac{m}{p} - \frac{n}{q}$$

since $dp = -dq$.

If the number of trials is zero, and we are in a state of complete ignorance with regard to the probability of success, we have

$$\ln \frac{p}{q} = 0$$

and since

$$q = 1 - p$$

then

$$\ln \frac{p}{1 - p} = 0$$

This yields $p = \frac{1}{2}$ on solution; that is, if we are in a state of no knowledge with regard to the probability of success, we must regard success or failure as equally likely.

In accordance with this maximizing principle, Fig. 1 and Table 1 have been prepared as graphic solutions of the measure-of-information equation. Figure 1 has as abscissa the total number of trials, N, which is to be read at integral values only. The curves are symmetrical about the probability ordinate of 0.5. The left ordinate gives the probability of success (reliability) for a given curve representing m successes and N - m failures; the right ordinate gives the probability of failure (unreliability) for n failures and N - n successes. Table 1 is the numerical equivalent of the graphs of Fig. 1.

When success probability is computed as the number of successes divided by the number of trials, the probability oscillates through extremes in the beginning and then converges slowly to the value for a large number of trials. The values presented in Fig. 1 and Table 1 start at the middle of the scale for zero trials, and as the results of more trials become available the oscillation is less extreme and the computed probability converges somewhat more rapidly to the value for a large number of trials.

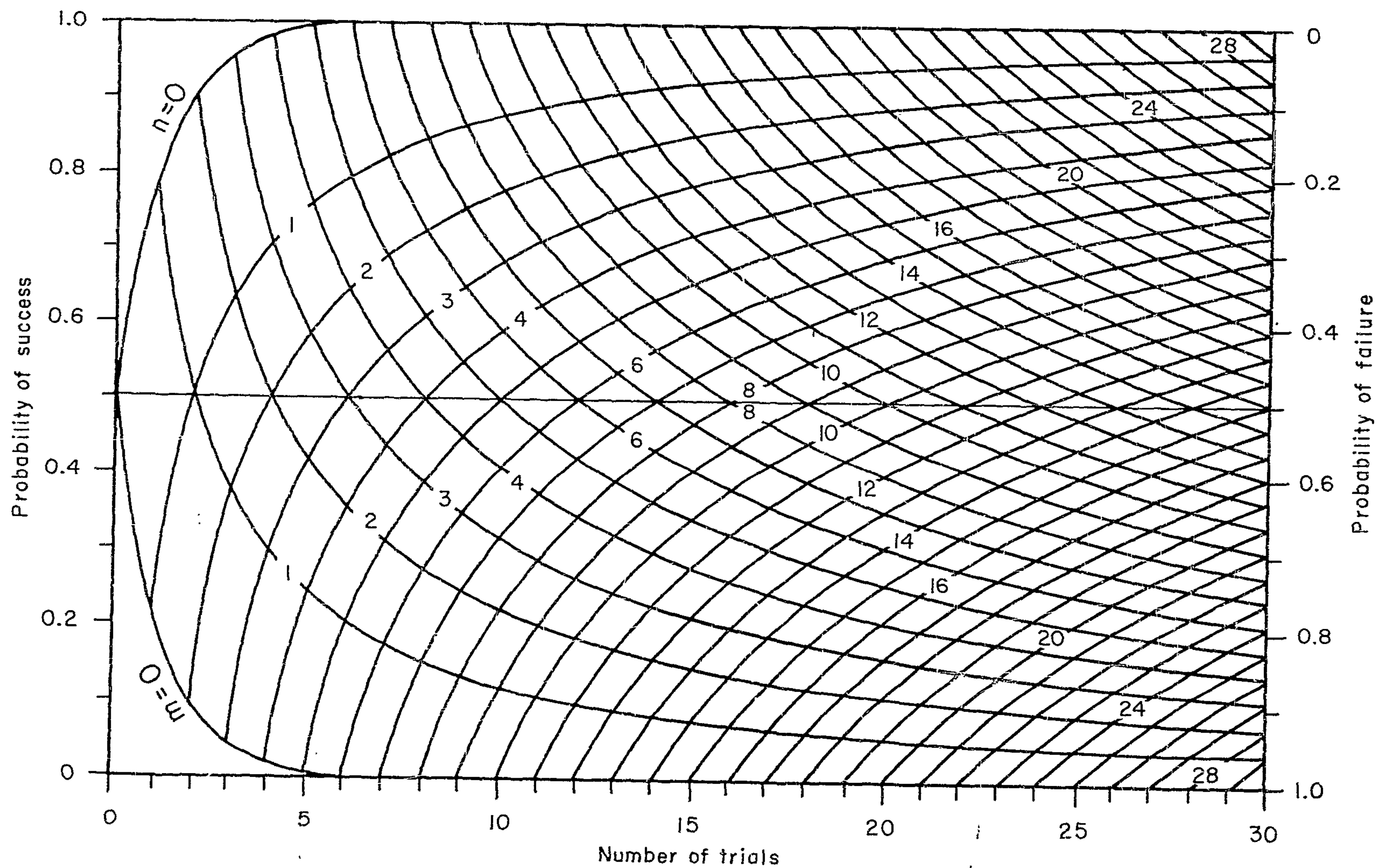

Fig. 1—Probability assignment in Bernoulli trials
(after having observed m successes and n failures previously)

Table 1

PROBABILITY OF SUCCESS FOR m SUCCESSES IN N TRIALS

m \ N	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
0	.500	.218	.098	.042	.017	.006	.002	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
1		.782	.500	.374	.295	.242	.204	.175	.153	.135	.121	.109	.100	.092	.085	.079	.073	.069	.065	.061	.058	.055	.052	.050	.047	.045	.043	.042	.040	.039	.037
2			.902	.626	.500	.416	.356	.310	.274	.245	.222	.202	.185	.171	.159	.148	.139	.130	.123	.116	.110	.105	.100	.095	.091	.087	.084	.081	.078	.075	.072
3				.958	.705	.584	.500	.437	.388	.349	.317	.289	.266	.247	.230	.215	.201	.190	.179	.170	.161	.153	.146	.140	.134	.129	.124	.119	.115	.111	.107
4					.983	.758	.644	.563	.500	.450	.409	.375	.345	.320	.298	.279	.262	.247	.234	.222	.211	.201	.192	.184	.176	.129	.162	.156	.151	.146	.141
5						.994	.796	.690	.612	.550	.500	.458	.423	.393	.366	.343	.323	.304	.288	.273	.260	.248	.237	.227	.217	.209	.201	.193	.187	.180	.174
6							.998	.825	.726	.651	.591	.542	.500	.464	.433	.406	.382	.361	.342	.324	.309	.294	.281	.269	.258	.248	.239	.230	.222	.214	.207
7								.999	.847	.755	.683	.625	.577	.536	.500	.464	.441	.417	.395	.376	.357	.340	.325	.312	.299	.287	.277	.267	.257	.249	.240
8									1.000	.865	.778	.711	.655	.607	.567	.536	.500	.472	.447	.425	.405	.386	.369	.354	.340	.326	.314	.303	.292	.282	.273
9										1.000	.879	.798	.734	.680	.634	.594	.559	.528	.500	.475	.452	.432	.413	.396	.380	.365	.352	.339	.327	.316	.306
10											1.000	.891	.815	.753	.702	.657	.618	.583	.553	.525	.500	.477	.457	.437	.420	.404	.389	.375	.362	.350	.338
11												1.000	.900	.829	.770	.721	.677	.639	.605	.575	.548	.523	.500	.479	.460	.442	.426	.411	.396	.383	.371
12													1.000	.908	.841	.785	.738	.696	.658	.624	.595	.568	.543	.521	.500	.481	.463	.446	.431	.417	.403
13														1.000	.915	.852	.799	.753	.712	.676	.643	.614	.587	.563	.540	.519	.500	.482	.466	.450	.435
14															1.000	.921	.861	.810	.766	.727	.691	.660	.631	.604	.580	.558	.537	.518	.500	.484	.468
15																1.000	.927	.870	.821	.778	.740	.706	.675	.646	.620	.596	.574	.554	.534	.516	.500
16																	1.000	.941	.877	.830	.789	.752	.719	.688	.660	.635	.611	.589	.569	.550	.532
17																		1.000	.935	.884	.839	.799	.763	.731	.701	.674	.648	.625	.604	.583	.565
18																			1.000	.939	.890	.847	.808	.773	.742	.713	.686	.661	.638	.617	.597
19																				1.000	.942	.895	.854	.816	.783	.752	.723	.697	.673	.650	.629
20																					1.000	.945	.900	.860	.824	.791	.761	.733	.708	.684	.662
21																						1.000	.948	.905	.866	.831	.799	.770	.743	.718	.694
22																							1.000	.950	.909	.871	.838	.807	.778	.751	.727
23																								1.000	.953	.913	.876	.844	.813	.786	.760
24																									1.000	.953	.916	.881	.849	.820	.793
25																										1.000	.957	.919	.885	.854	.826
26																											1.000	.958	.922	.889	.859
27																												1.000	.960	.925	.893
28																													1.000	.961	.938
29																														1.000	.973
30																															1.000

III. EXAMPLES

Figure 2 presents graphically the reliability results, based on Ref. 7, of the first 15 Atlas missile firings. The solid curve is computed by the method of maximizing the uncertainty. The dashed curve is based on the conventional method of number of successes divided by number of trials. In the beginning the solid curve oscillates less violently than the dashed curve. As the number of trials grows larger, the curves approach each other. The relative initial stability of the solid curve will lead to a reliability less subject to rapid change as successive initial firings occur.

Figure 3 is a plot, similar to Fig. 2, obtained by using random numbers to simulate repeated tossing of a perfect coin. Even numbers were assigned to "success" and odd numbers to "failure." The solid curve again shows less oscillation initially than the dashed curve. The two curves approach each other as the number of trials increases.

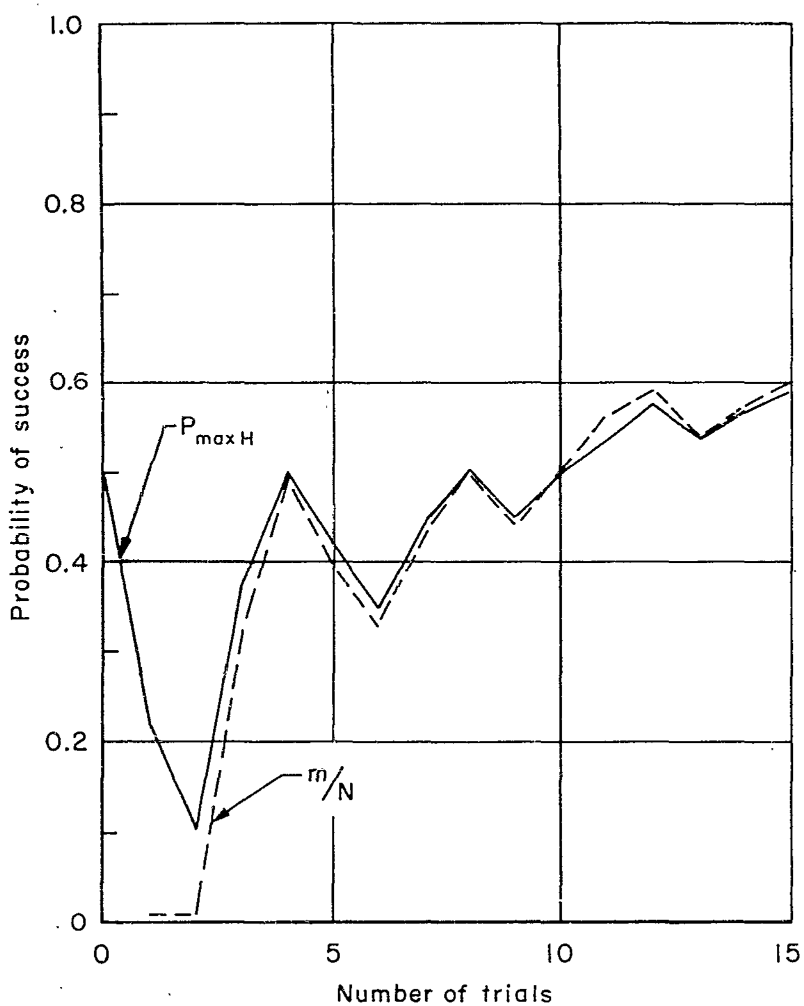

Fig. 2 — Reliability results of Atlas firings

No.	Success	Failure	$P_{max\,H}$	$^m/_N$
0			0.50	---
1		X	0.22	0
2		X	0.10	0
3	X		0.37	0.33
4	X		0.50	0.50
5		X	0.42	0.40
6		X	0.35	0.33
7	X		0.44	0.43
8	X		0.50	0.50
9		X	0.45	0.44
10	X		0.50	0.50
11	X		0.54	0.56
12	X		0.58	0.59
13		X	0.54	0.54
14	X		0.57	0.57
15	X		0 59	0.60

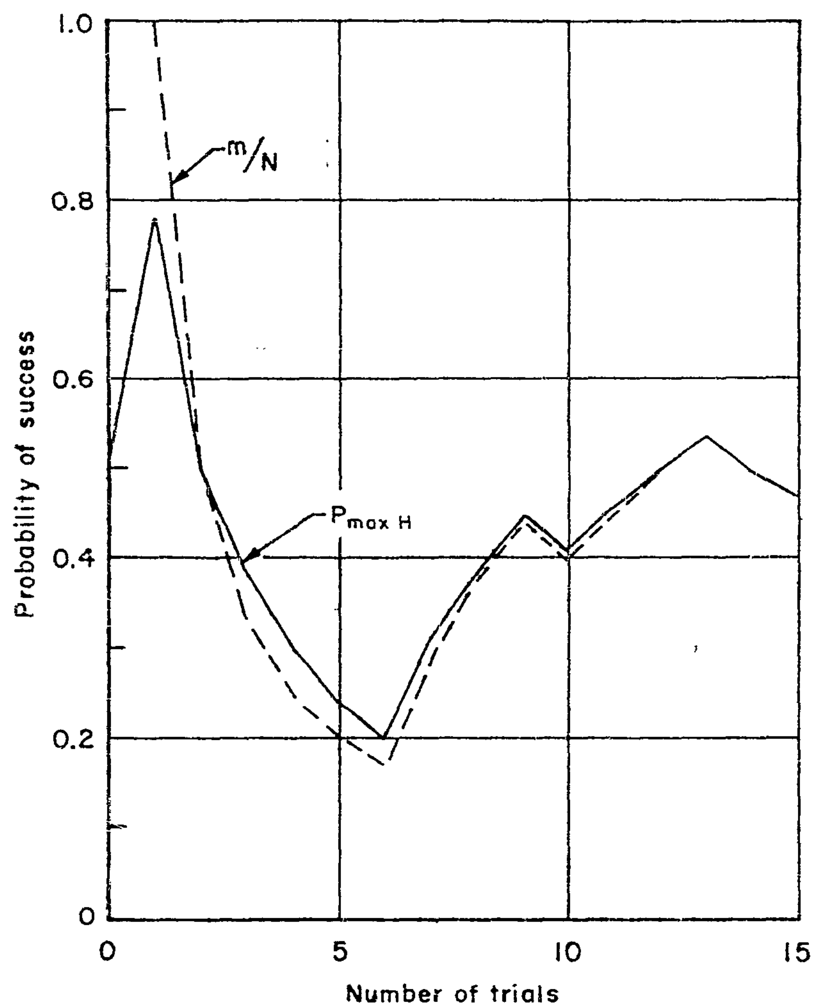

Fig. 3 — Reliability results using random numbers to simulate tossing of perfect coin

No.	Success (even)	Failure (odd)	$P_{max\,H}$	m/N
0	--	--	0.50	--
1	X		0.78	1.00
2		X	0.50	0.50
3		X	0.38	0.33
4		X	0.30	0.25
5		X	0.24	0.20
6		X	0.20	0.17
7	X		0.31	0.28
8	X		0.39	0.38
9	X		0.45	0.44
10		X	0.41	0.40
11	X		0.46	0.45
12	X		0.50	0.50
13	X		0.54	0.54
14		X	0.50	0.50
15		X	0.47	0.47

IV. CONCLUSIONS

The new method of point estimation of probabilities for success is characterized by less-extreme fluctuations as the results of Bernoulli trials become known. The estimates approach the conventional estimate of the ratio of successes to total trials as the number of trials becomes large. Application of the method to reliability testing should be made to determine if predictions of reliability can be made with more assurance for a given number of units tested when sample sizes are very small.

Further areas of interest include application of the method to problems where there is some knowledge of reliability before testing, and development of interval estimates for various confidence levels based on the binomial distribution. Comparison of the properties of this distribution with those suggested by Steinhaus(1) should also be instructive.

REFERENCES

1. Steinhaus, H., "The Problem of Estimation," Ann. Math. Stat., Vol. 28, No. 3, September 1957.

2. Shannon, C. E., and W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, Urbana, Illinois, 1949.

3. Woodward, P. M., Probability and Information Theory, with Applications to Radar, McGraw-Hill Book Co., Inc., New York, 1953.

4. Jaynes, E. T., "Information Theory and Statistical Mechanics," Phys. Rev., 2nd Series, Vol. 106, No. 4, May 15, 1957.

5. Jaynes, E. T., "Information Theory and Statistical Mechanics, II," Phys. Rev., 2nd Series, Vol. 108, No. 3, October 15, 1957.

6. Tribus, Myron, Thermostatics and Thermodynamics, D. Van Nostrand Co., Inc., Princeton, New Jersey, 1961.

7. A Chronology of Missile and Astronautic Events, 87th Cong., 1st sess., House Report No. 67, Union Calendar 34, U. S. Government Printing Office, Washington 25, D. C., March 8, 1961.

UNCLASSIFIED

UNCLASSIFIED